RESEARCH ARTICLE

Spatial dynamics of spotted lanternfly, *Lycorma delicatula*, invasion of the Northeastern United States

Rachel T. Cook[1], Samuel F. Ward[2], Andrew M. Liebhold[3,4], Songlin Fei[1]

1 *Purdue University Department of Forestry and Natural Resources, 715 West State Street, West Lafayette, IN, USA* **2** *Mississippi State University Department of Biochemistry, Molecular Biology, Entomology and Plant Pathology, 100 Old Highway 12, Clay Lyle Entomology Building, Mississippi State, MS, USA* **3** *USDA Forest Service Northern Research Station, 180 Canfield Street, Morgantown, WV, USA* **4** *Faculty of Forestry and Wood Sciences, Czech University of Life Sciences Prague, 165 00, Praha 6 – Suchdol, Czech Republic*

Corresponding author: Songlin Fei (sfei@purdue.edu)

Abstract

Spotted lanternfly (SLF), *Lycorma delicatula* (White) (Hemiptera: Fulgoridae), is a non-native planthopper that recently established in the Northeastern United States. Little is known about the spatial dynamics of its invasion and key drivers associated with its regional spread. Here, using field survey data from a total of 241,366 survey locations from 2014–2019 in the eastern USA, we quantified rates of SLF spread and modeled factors associated with the risk of SLF invasion. During the study period, SLF invasion appears to be associated with both short- and long-distance dispersal. On average, the number of newly invaded counties per year increased since initial discovery, with 0–14 long-distance dispersal events per year and median jump distances ranging from 55 to 92 km/year throughout the study period. Radial rates of spread, based on two of the three analysis methods applied, varied from 38.6 to 46.2 km/year. A Cox proportional hazards model suggested that risk of SLF invasion increased with a proxy for human-aided dispersal, human population per county. We anticipate that SLF will continue to spread via both long- and short-distance dispersals, especially via human activities. Efforts to manage SLF populations potentially could target human-mediated movement of SLF to reduce rates of spread.

Keywords

Biological invasion, Cox proportional hazards, spatiotemporal, invasive species, radial spread

Introduction

Though most non-native pests fail to establish after arrival, those that successfully found reproducing populations can subsequently spread via a coupling of population growth with dispersal. The dispersal of many invading species is characterized by the simultaneous occurrence of local diffusion and occasional long-distance dispersal (Hastings 2005; Liebhold and Tobin 2008). Information on what factors drive spread of a non-native pest can guide management to contain its populations and reduce their impacts to ecosystems and economic costs (Sharov and Liebhold 1998; Liebhold and Kean 2019). Understanding the factors that drive spread is particularly important for newly established species, for which dispersal behaviors and population growth characteristics are often unknown.

Spotted lanternfly, *Lycorma delicatula* (White) (Hemiptera: Fulgoridae), is a non-native planthopper recently established in the United States. The species is native to southeast Asia, but recently invaded the USA in Berks County, Pennsylvania in 2014 (Barringer et al. 2015). Spotted lanternfly (SLF) is univoltine and lays egg masses on a variety of surfaces, including tree bark, stone, motor vehicles, and trains (Urban 2019). In addition to indiscriminate egg deposition, SLF also has a wide breadth of host use. This pest feeds on over 70 species of herbaceous and woody plants belonging to over 20 families, though it prefers tree of heaven (*Ailanthus altissima*), especially as a late instar (Dara et al. 2015; Parra et al. 2017). Notably, SLF feeds on apple (*Malus* spp.) and grape (*Vitis* spp.), both important agricultural plants in the Northeastern USA. Feeding on grape has reportedly resulted in lower fruit quality, less fruit production, and elevated mortality, though minimal impacts to fruit tree health have been reported (Urban 2019). The most conspicuous impact of SLF in forests is the accumulation of honeydew in the understory, which results in sooty mold growth that limits photosynthesis and growth of understory plants (Ding et al. 2006; Parra et al. 2017). There is also evidence that aggregation of SLF can cause weeping wounds on trees, resulting in crown dieback (Dara et al. 2015). While detrimental impacts on tree of heaven might be beneficial due to its status as an invasive plant, SLF is considered a serious pest due to its negative impacts on agricultural crops and native trees.

Despite regulations by the state of Pennsylvania that prohibit movement of any SLF living stage (e.g. egg masses, nymphs, adults) or material potentially harboring the pest (e.g. firewood, nursery stock, etc.) outside of a quarantine area, SLF has spread from Pennsylvania to seven surrounding states as of 2019 (Fig. 1). Because SLF was detected in the USA recently, there is little information on how this species spreads or what drives its invasion. Though the body of knowledge on this insect is growing, many aspects of SLF spread, especially the role of environmental drivers, are unknown. Elucidating how this pest spreads can inform future management and survey efforts.

The ranges of introduced species are influenced by a multitude of anthropogenic factors and habitat features. For SLF, climatic niche models indicate that half of the USA, including most of the New England, Mid-Atlantic, and Pacific Coast states, is at risk of invasion (Wakie et al. 2020). While these climatic niche models provide valu-
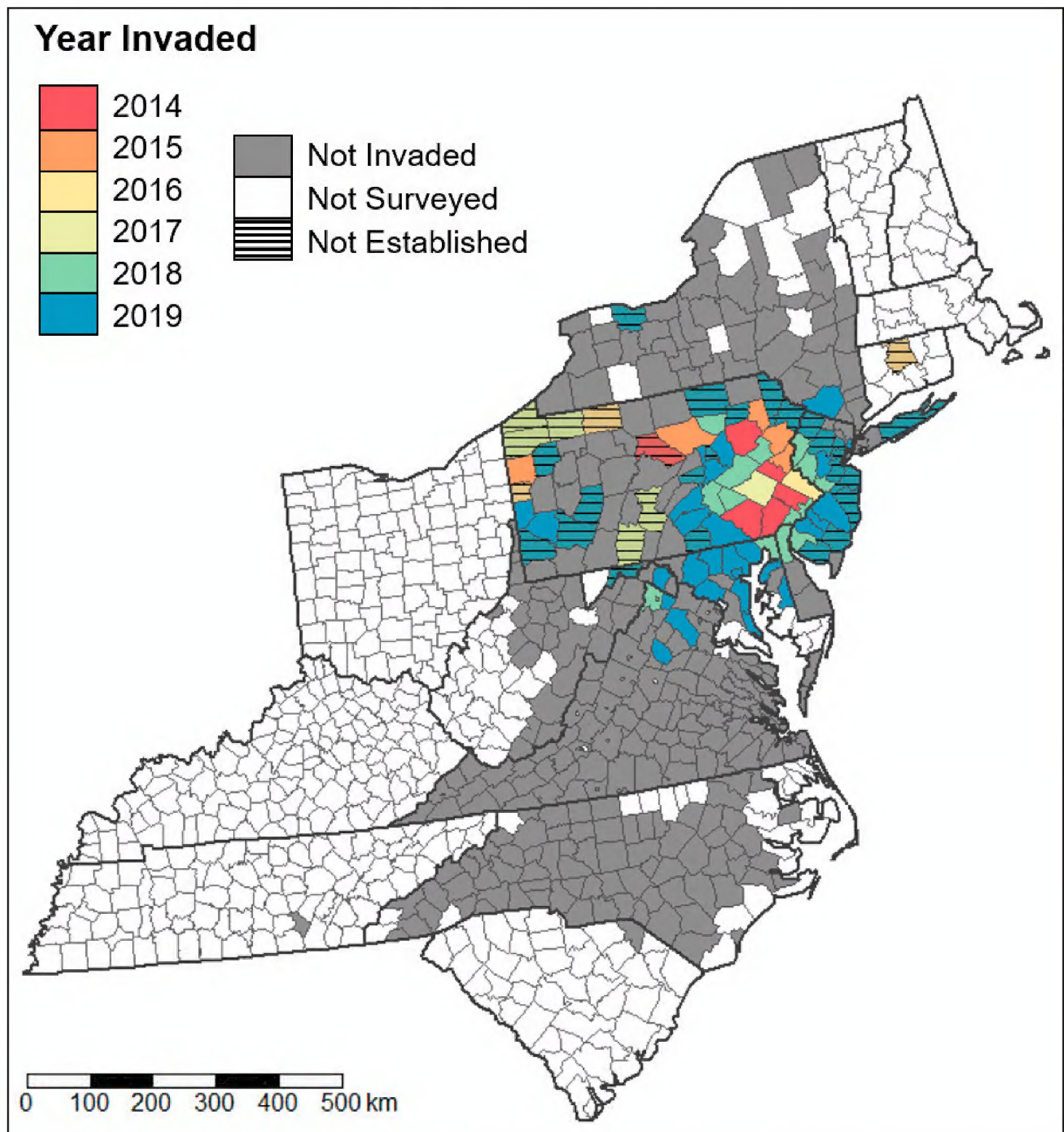
Figure 1. County-level distributions of spotted lanternfly (SLF) in the eastern USA. Distribution of SLF detections and establishments by year based on USDA Animal and Plant Health Inspection Service and Pennsylvania Department of Agriculture visual survey data. Counties with hash marks had SLF detections that failed to establish. We define a county as invaded when the county experiences at least two consecutive years of SLF detection, and define year invaded as the first of those two consecutive years. Counties with white color were not surveyed.

able information on where SLF can potentially establish, analyses of spread can provide insight into how quickly SLF will arrive and what habitat and/or anthropogenic factors affect the dynamics of SLF spread. SLF has undergone several long-distance dispersal events that likely resulted from human-mediated transportation (Eddy 2018; Scheid 2020). Tree of heaven is also more abundant in urban areas, and thus human activities may increase both propagule pressure and habitat suitability. However, the rate

of spread, including the frequency and distance of long-distance dispersal events, and drivers of spread have not been quantified. Therefore, we investigated how anthropogenic and habitat factors are related to SLF spread.

We analyzed the known geographical distribution of SLF (2014–2020) in the USA to quantify its rates of spread and identify factors that influence its invasion risk. Our goals were to: 1) describe the patterns of SLF spread following the initial detection in 2014, and 2) identify key drivers that are associated with SLF spread. For our second goal, we used known occurrences of SLF in conjunction with habitat and anthropogenic variables to determine the most important factors driving county-level invasion risk across the study area, defined below. We hypothesized that anthropogenic factors are important drivers of SLF spread, given the ability of this insect to lay inconspicuous eggs on a variety of materials, including motor vehicles and trains (Urban 2019).

Methods

The SLF distribution data analyzed in this study were derived from visual surveys conducted from 2014–2019 by the US Department of Agriculture, Animal and Plant Health Inspection Service (APHIS) and the Pennsylvania Department of Agriculture (PDA). We also used SLF county-level presence data for 2020 from the New York State Integrated Pest Management Program (Cornell 2021). Survey data include geospatial coordinates for survey locations as well as the number of SLF observed (if any). A total of 241,366 survey locations were obtained for this study (Suppl. material 1). Given the irregularity of survey locations and potential biases (e.g. surveys at expected SLF locations) and to render data at an equivalent scale as the 2020 presence data, we converted counts to county-level presence/absence records and used county as our unit of analysis.

The survey data contained many points that we identified as failed establishments in which SLF were observed in a county in a given year but were absent in surveys of the same county in subsequent years. These detections were likely either populations that failed to establish or regulatory incidents, such as dead SLF adults found in transported materials, and thus we did not treat them as invasions. Hereafter, we refer to detections as establishments plus failed establishments and establishments as only populations that persisted for more than one survey year in consecutive years within a county. Moreover, we categorized each invaded county in year n as contiguous or non-noncontiguous based on the presence or absence, respectively, of an invaded neighboring county in year n-1.

Described below are methods we used to 1) determine aspects of spread dynamics, such as jump distances and spread events into contiguous vs. non-contiguous counties, 2) compare three methods of estimating spread rates, and 3) fit a Cox proportional hazards model estimating time-to-invasion as a function of variables representing spatial proximity to existing SLF populations (henceforth referred to as spatial proximity), habitat suitability, and anthropogenic influences. Our study area was defined as the area of the eastern USA invaded in 2019 plus a buffer distance of 355 km, equal to the maximum observed jump distance (see "Characterization of spread events" in

Methods). This study area was used for all subsequent analyses. Counties, which are the level at which quarantines and other management decisions are set, served as the unit of analysis for all analyses. All analyses were conducted using R version 4.0.2 (R Core Team 2020).

Characterization of spread events

To characterize spread, we quantified the number of yearly spread events into contiguous and non-contiguous counties, as well as the distribution of jump distances. Jump distance is defined as the distance between establishments or detections in non-contiguous counties in year n and the nearest previously invaded county in year n-1. We estimated jump distances for every newly invaded county by calculating the distance to the closest previously invaded county, as assuming new SLF establishments originate from the closest previously invaded county provides a conservative estimate. Distances were measured using county centroids. We repeated this process for each year, and summarized the distribution of jump distances (e.g. median, minimum, maximum). To determine if spatial proximity is related to whether or not a detection became an establishment (i.e. an invasion persisted), we separated jump distances by establishments and failed establishments and used a Mann-Whitney U test to compare the distribution of jump distances between these two groups.

Spread rates

Because little is currently known about SLF spread patterns and different approaches can provide variable estimates of annual spread (Tobin et al. 2006), we compared three methods to calculate spread rate described by Gilbert and Liebhold (2010). The purpose of our comparison of these methods is to provide a range of possible spread rates as well as to determine robustness of each when applied to an insect at early stages of invasion.

The first method is to apply regression of the distance (centroid to centroid Euclidean distance) of every county with positive establishment from the point of initial detection (Berks County, PA) as a function of years since initial detection (2014). The resulting slope of the estimated regression equation estimates the radial rate of spread measured in distance/year. The second method is to regress the square root of the invaded area (estimated by summing the area of invaded counties in each year) divided by π on time. The resulting slope of the estimated regression line estimates the radial spread rate in distance/year (e.g. effective range radius; Shigesada et al. 1995). Last, we calculated the average distance between invasion boundaries in consecutive years along radii emanating every 0.5 degrees from the centroid of Berks County, PA. We used radii at a frequent degree interval to obtain a high-resolution estimate of yearly distance between boundaries. We found invasion boundaries by fitting a convex hull polygon to the area of invasion in each year and subsequently converting the polygon edges to lines. The convex hull polygon in each year was stretched to the edges of non-contiguous invaded counties. The resulting average distance between boundaries on each radius between consecutive years can be used to estimate the annual radial spread rate (e.g. boundary

displacement rate). Due to the nature of fitting a convex hull polygon around invaded counties, we used county boundaries as opposed to county centroids to calculate distances in boundary displacement estimations. In summary, distance regression is based on distance and year of sampling points from the origin where the species was first detected, while effective range radius considers area invaded over time. Boundary displacement estimates distance between invasion boundaries in consecutive years.

Dispersal kernel estimation

Dispersal kernels estimating risk of invasion as a function of distance have been developed for other invading forest insects (Orlova-Bienkowskaja and Bienkowski 2018). Given interspecific variation in spread rates (Fahrner and Aukema 2018), however, we estimated a SLF specific dispersal kernel, which, in turn, should enable more reliable estimates of the effects of SLF spatial proximity on invasion risk. Our analysis used 2015–2019 county-level SLF survey data from USDA APHIS and PDA and follows methods from Kovacs et al. (2010). A negative exponential function was used to model the probability, p, of each non-invaded county in the study area becoming invaded on an annual basis from 2015–2019:

$$p_{i,j} = e^{-ad} \quad (1)$$

where a is the parameter we sought to estimate and d is the distance in kilometers to a previously invaded county. To estimate a, we simulated county-level spread starting from the five initially invaded counties in 2014 using values of a between 0.01 and 0.10 in 0.001 intervals.

To simulate spread for a given a value, we calculated the centroid to centroid Euclidean distance from each non-invaded county i in year n to each invaded county j as of year n-1, as each county j invaded as of year n-1 could serve as a source for invasion into county i in year n. The distances from county i to each invaded county j were input into Equation 1, producing an estimate, p, for the probability of SLF invading from each county j. This probability value was then used to parameterize a Bernoulli distribution such that the probability of an event was equal to p. We then took a random draw from that Bernoulli distribution in which a draw of 1 or 0 would indicate invasion or non-invasion, respectively. This meant that there were x draws for each non-invaded county i, where x = number of invaded counties in year n-1. If any draw produced a 1, the county was categorized as invaded for the rest of the simulation (i.e. counties could not become uninvaded).

A single iteration of this process produced a simulated, county-level invasion at annual time steps (2015–2019) that may or may not have reflected the realized invasion. For each a value, we conducted 500 iterative simulations, starting with the initially invaded counties in 2014 and forecasting spread to 2019. Results were summarized with accuracy values - false negatives and positives, and true negatives and positives - compared with the actual invasion data from 2015–2019. We selected the value of a that simultaneously resulted in the lowest number of false negatives and false positives when comparing actual spread to predicted spread.

Invasion drivers

Cox proportional hazards models can be used to estimate survival time based on predictor variables, including both static and time-varying predictors (Thomas and Reyes 2014). If we equate survival to a county persisting without invasion, we can use Cox proportional hazards models to evaluate which factors explain variation in time-to-invasion. Therefore, we used a Cox proportional hazards time-to-invasion model to evaluate potential drivers of SLF invasion at the county level, in a manner analogous to the implementation by Jules et al. (2002) and Ward et al. (2020).

The Cox proportional hazards model quantifies the probability of invasion at each one-year time step. Time steps ranged from 2014–2015 to 2018–2019. Predictor variables included static habitat variables (Suppl. material 4) and one time-varying predictor, spatial proximity. To quantify spatial proximity, we first used Equation 1, setting $a = 0.045$ (i.e. determined from the dispersal kernel estimation process described above; see "Invasion drivers" in Results) and d as the distance in kilometers between each uninvaded county to all previously invaded counties. Spatial proximity, denoted by *SpatialProx*, was then calculated for each county:

$$SpatialProx_i = 1 - \prod(1 - p_{i,j}). \quad (2)$$

The other predictors included two anthropogenic variables and six habitat variables. The anthropogenic variables were human population from the U.S. Census and road density calculated by Liebhold et al. 2013 from the ArcGIS World Transportation reference layer (Suppl. material 2) and each was considered a proxy for human-aided dispersal. The six habitat variables included forested area and five host availability terms expressed as basal area, host trees per acre, number of host trees per county, tree of heaven occurrence, and canopy cover (Suppl. material 3). Forested land was obtained from the US Forest Service FIA MapMaker online data query system (https://www.nrs.fs.fed.us/fia/data-tools/mapping-tools). Percent forest canopy cover was obtained from the Forest Service's cartographic tree canopy cover product (USDA Forest Service 2016).

Host basal area and numbers of host trees per acre and county were obtained from the Forest Service's Forest Inventory and Analysis (FIA) program, using a published list of known SLF hosts from Barringer and Ciafre (2020). The FIA program is a long-term forest inventory program with one 0.40 hectare sample every 2,428 hectares, with most counties partially assessed annually since 2000. FIA assesses forest areas defined as at least 37 meters wide and 0.40 hectares in size, covered by at least 10% trees (Bechtold and Patterson 2005). We obtained plot-level basal area and stem density per acre from FIA records from 2015–2017. To estimate these variables at the county-level, we aggregated each by the summed county plot area for every known host with available FIA data by species code, obtained from the National Core Field Guide (USDA Forest Service 2019). We then estimated the number of each host species in a county by multiplying the estimated number of trees per acre by the total acres of forested land in each county. Because FIA only surveys forested areas, and tree of heaven is often found in developed or urbanized areas, a number of tree of heaven observations were down-

loaded separately as point data from EDDMapS (EDDMapS 2021) and aggregated to the county level by summing up the number of observations per county.

Prior to model development, we quantified pairwise correlations between our predictors to check for collinearity (defined as Pearson's product moment correlation coefficient ≥ 0.70). Based on this step, we removed road density and number of host trees per county due to collinearity with human population and forested area, respectively. We removed these two variables as opposed to human population and forested area because in preliminary models, they were more strongly associated (i.e. occurred in models with lower Akaike Information Criterion values) with SLF time-to-invasion than their co-varying counterparts. We then refined the model by applying a backward selection procedure that iteratively removed the variable associated with the highest p-value and refitting the model until only statistically significant predictors remained.

Results

Characterization of spread events

There was overall an upward trend in the number of newly invaded counties every year since initial discovery, although some counties contained failed establishments. There was a drop in counties with establishments in 2016 and 2017, while the highest number of establishments was observed in 2019 (Table 1). Similar to the number of newly invaded counties per year, number of counties with failed establishments generally increased across the study period and peaked in 2019. The highest percent of counties with failed establishments occurred in 2017, with 83% of detections failing to establish. The median yearly jump length into counties with detection and establishment ranged from 46 to 73 km and 50 to 92 km, respectively.

We did not find a significant difference between distributions of jump distances in established populations vs. failed establishments. Median jump distances across all years in failed establishments and established populations were 55 km and 71 km, respectively. A Mann-Whitney U test showed the distributions in the two groups did not significantly differ (W = 706, p = 0.46).

Current data suggest that the SLF invasion began in eastern Pennsylvania, and many of the counties invaded in the surrounding area of eastern and central Pennsylvania were contiguous with previously invaded counties (Fig. 2). In contrast, several populations in western Pennsylvania and northern Virginia resulted from invasion into non-contiguous counties, indicating long-distance jumps. There were no newly invaded non-contiguous counties in 2016 or 2017. Trends in the number of both contiguous and non-contiguous counties tracked the overall number of counties invaded, starting out low and increasing in 2018 and again in 2019 (Table 1). However, there were overall fewer non-contiguous counties invaded than contiguous counties across all years.

Establishments showed similar patterns in numbers of new counties invaded and jump distances by year (Table 1), with lower values from 2014 to 2017 and an increase

Table 1. Spread events summary. Number of observed contiguous (having at least one previously invaded neighboring county at time of invasion) and non-contiguous (having no previously invaded neighboring counties at time of invasion) newly invaded counties per year and median jump distances between invaded and uninvaded counties between consecutive years for both the non-persistent and the persistent counties.

	2014	2015	2016	2017	2018	2019	Total
Counties with detections (n)	5	6	4	6	18	47	86
Counties with establishment (n)	5	5	1	1	15	27	54
Counties with failed establishments (n)	0	1	3	5	3	20	32
% of Counties with failed establishments	-	16.7	75.0	83.3	16.7	42.6	-
Median jump length (km) into counties with detection	-	137.4	100.5	79.6	104.5	46.5	-
Median jump length (km) into counties with establishment	-	54.5	49.9	69.5	91.7	57.8	-
Counties with contiguous invasion	5	3	1	1	12	13	35
Counties with non-contiguous invasion	-	2	0	0	3	14	19

in 2018 and 2019. Median jump distances were greatest in 2017–2018 in the established counties, and were greatest in 2014–2015 in counties with detections. The year with the highest number of counties with newly established populations was 2019, whereas the year with the largest median jump distance (92 km) was 2018. Median jump distances were generally higher into counties with detections than counties with establishments. The overall maximum jump distance was 355 km into Mercer County in northwest Pennsylvania (Fig. 3), while the median jump distance was 55 km for detections and 71 km for establishments.

Spread rates

Estimated spread rates varied from 15–46 km per year among our three methods. Spread rate estimated by effective range radius was 46.2 km/year (SE = 7.19 km, 95% CI 26.26-66.20; Fig. 4A). Spread rate was estimated at 15.2 km/year (SE = 6.40 km, 95% CI 2.35-28.03) using distance regression (Fig. 4B). Spread rate estimated by average boundary displacement (averaged over all years) was 38.6 km/year (range 0 to 75 km; Fig. 4C), which was approximately 10 km less than estimated by effective range radius. The median boundary displacement across all years was 20.8 km/year. There was no difference in invaded area boundaries between 2016 and 2017, because the only newly invaded county was within the existing invasion boundary.

Invasion drivers

The best fitting value of a in the exponential dispersal kernel (*Equation* 1) was 0.045, which simultaneously resulted in the lowest number of false negatives and false positives. We used this value to estimate spatial proximity in the Cox proportional hazards model.

In the final Cox proportional hazards model, the hazards ratios for both spatial proximity and human population were greater than 1, indicating a positive relationship with increased risk of invasion (Table 2). Spatial proximity was identified as the strongest predictor (i.e. highest Z-value) with a notably high hazards ratio of ~40, followed by human population. No other covariates were statistically significant.
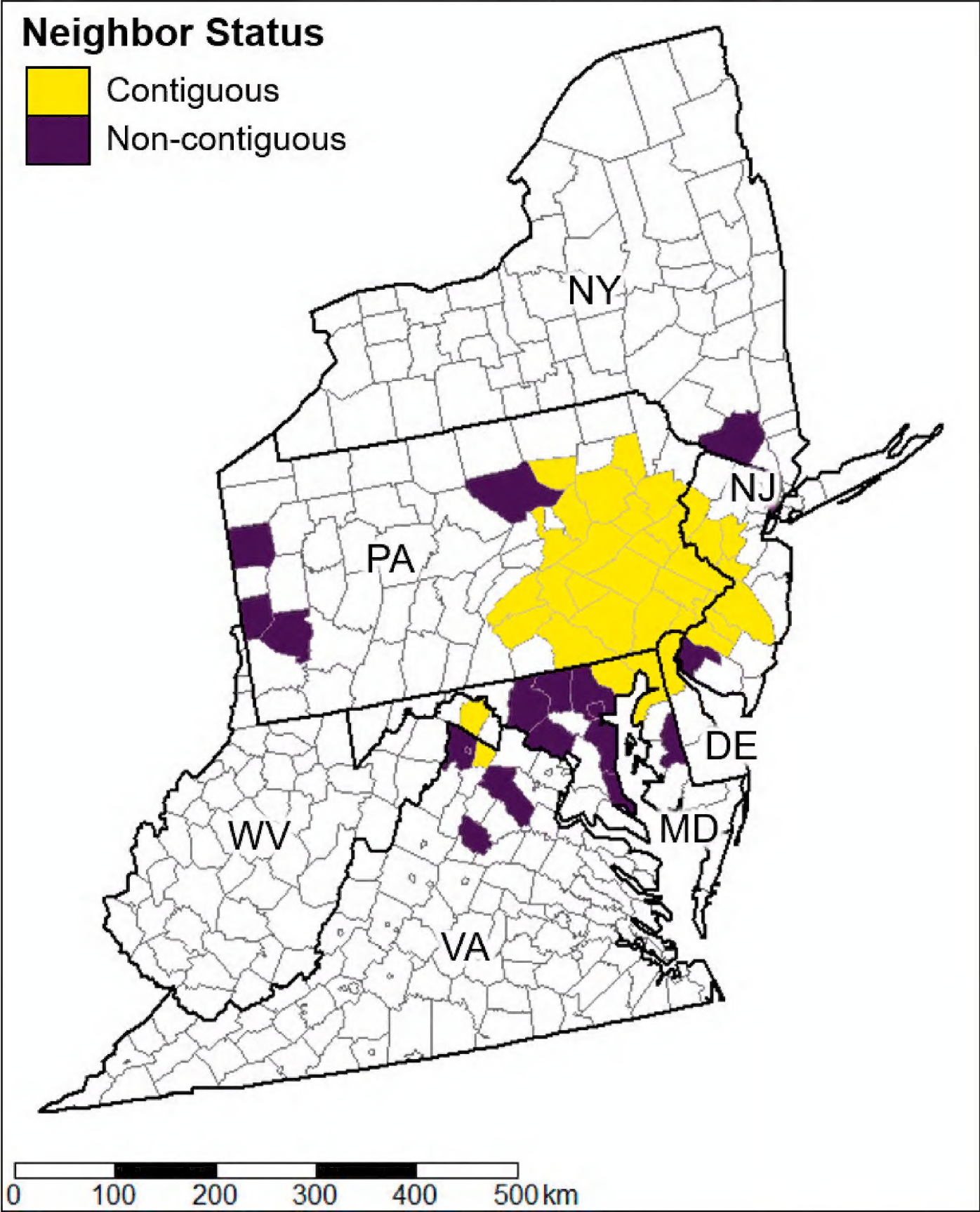
Figure 2. Contiguous and non-contiguous establishments of spotted lanternfly. Spatial distribution of contiguous (having at least one previously invaded neighboring county at time of invasion) and non-contiguous (having no previously invaded neighboring counties at time of invasion) counties across the study area.

Table 2. Final Cox proportional hazards model summary. Summary statistics from final Cox proportional hazards model predicting time-to-invasion of SLF at the county level in the study area.

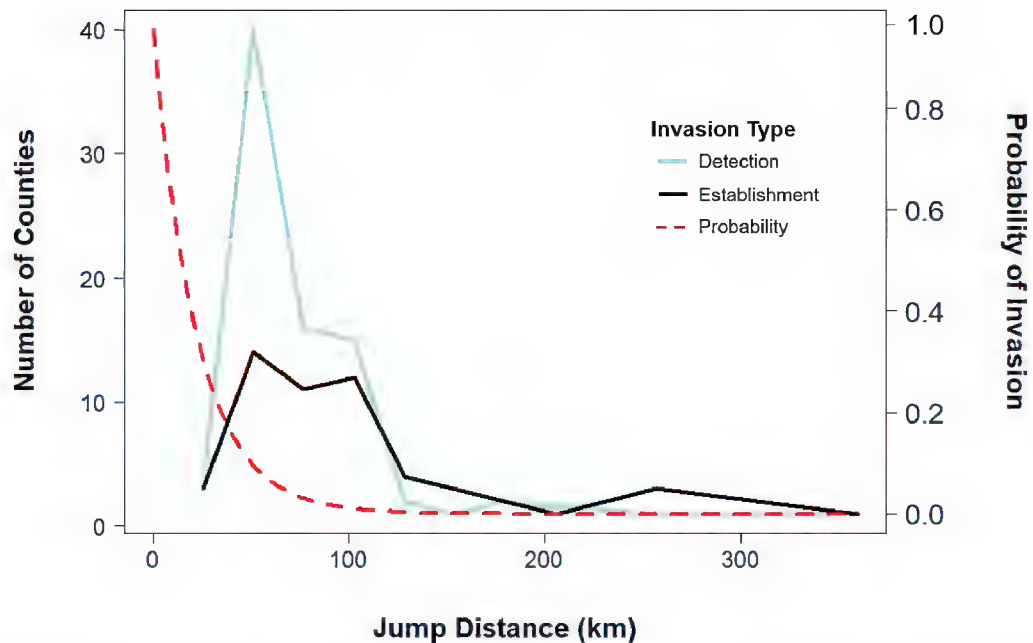| Predictor	Estimate (coefficient)	SE	Z	p-value	Hazards ratio (95% CI)
SLF spatial proximity	3.70	0.286	12.94	<0.0001	40.29 (23.01-70.54)
Human population	0.28	0.126	2.22	0.0265	1.32 (1.03-1.69)

Figure 3. Jump distance distributions and probability of invasion by spotted lanternfly (SLF). Line graph of observed jump distances (the distance between new establishments in year n and the nearest previously invaded county in year n-1) for every newly invaded county for both establishments (black) and detections (blue). The red line indicates the probability of invasion by distance, based on the estimated SLF-specific negative exponential kernel function $p_{ij} = e^{-0.045d}$.

Discussion

Spread of invasive species is often characterized by both short- and long-distance dispersal. In many systems, short-distance dispersal is caused by the natural movement of organisms (e.g. flight behavior) while long-distance dispersal is caused by accidental human movement (Hastings et al. 2005). Even small amounts of long-distance dispersal can result in greatly elevated rates of spread (Shigesada et al. 1995). So far in the SLF invasion, movement appears to consist of both short- and long-distance dispersal. Little is known about natural dispersal in this species. Our results suggest, however, that risk of long-distance movement increases with human population density, likely reflecting the propensity of SLF to become associated with objects transported by humans, such as when SLFs oviposit onto train cars and motor vehicles (Urban 2019).

A higher number of new establishments occurred in contiguous than in non-contiguous counties, but several long-distance jumps were observed and the frequency of jumps appears to be increasing (Table 1). Human-mediated long-distance dispersal events are responsible for spread outside of the center of invasion, allowing for invasion of a larger geographic area than would be possible via insect movement alone. For example, the established population in northern Virginia (Frederick County) is believed to have originated from shipments from a stone yard in Pennsylvania (Eddy 2018). As SLF spreads, there may be increases in both long- and short-distance movement due to increases in numbers of source populations or increases in population size. Indeed, there have been additional long-distance dispersals beyond our study period (2014–2019), such as the discovery of SLF in Switzerland County, IN in July 2021 and the identification of SLF in a bug collection at the Kansas State Fair in early September

2021 (Edwards 2021; Indiana DNR 2021). A Mann Whitney U test showed no significant difference between jump distance distributions in detected but non-established vs. established populations, indicating that jump dispersal events are not necessarily more likely to persist if they are closer to the point of establishment (i.e. have closer spatial proximity). Ranges of jump distances were visually similar in range for both detected and established populations (Fig. 3), signifying that established jumps went at least as far as jumps that failed to establish. Shigesada et al. (1995) demonstrated that such long-distance dispersal events typically result in faster rates of spread as well as accelerating patterns of radial spread. SLF spread rates could increase in this way, and we observed the largest increases in radial spread in the last two years of the study period (2018 and 2019), potentially indicating accelerating spread.

Our estimates of spread rate varied between methods, with the effective range radius method estimating the highest spread rate. The large differences observed between these methods may reflect the discontinuous nature of SLF spread. Measurement of the radial rate of spread of invading organisms was originally envisioned for continuous range expansion (e.g. Skellam 1951) and may not fully capture discontinuous spread such as observed here, which is also reflected in the low variance explained by distance regression spread estimation ($r^2 = 0.098$) (Fig. 4B). The effective range radius approach may provide a more representative measure of spread in this situation as it accounts for both the frequent long-distance dispersals and subsequent spread into the counties between contiguous and non-contiguous counties. For example, a long-distance dispersal event established a SLF population in northern Virginia in 2018, and in 2019, SLF spread to several counties between the eastern Pennsylvania invasion area and the new area in northern Virginia (Fig. 1). The effective range radius approach accounts for the cumulative invaded area as these counties are occupied in subsequent years, whereas boundary displacement does not include those counties in estimates of radial spread. That is, counties closer to the previously invaded area following a long-distance jump are enclosed by the convex hull polygon and do not influence future boundary displacements as they become invaded. However, the effective range radius approach alone may overestimate spread as entire county areas are summed as invaded, while only portions of each county are actually invaded.

Therefore, based on the findings presented here, we estimate the radial spread rate at around 40 km/year based on the average of the two more reliable methods (i.e. effective range radius and boundary displacement). If SLF were allowed to spread without any intervention, spread might be much higher given considerable management efforts are currently targeted to suppress SLF populations and limit movement. For example, active management programs conducted by USDA APHIS include egg scraping, sanitation (i.e. host tree removal) around SLF detections, and insecticide application to tree of heaven designated as trap trees (USDA-APHIS 2018). Insecticide applications were used primarily to kill SLF landing on trap trees and, in later applications, to determine efficacy of insecticides for use on fruit and residential trees (Urban, Calvin, and Hills-Stevenson 2021). Additionally, the State of Pennsylvania's quarantine on movement of goods out of the invaded area is implemented to limit spread of SLF. It
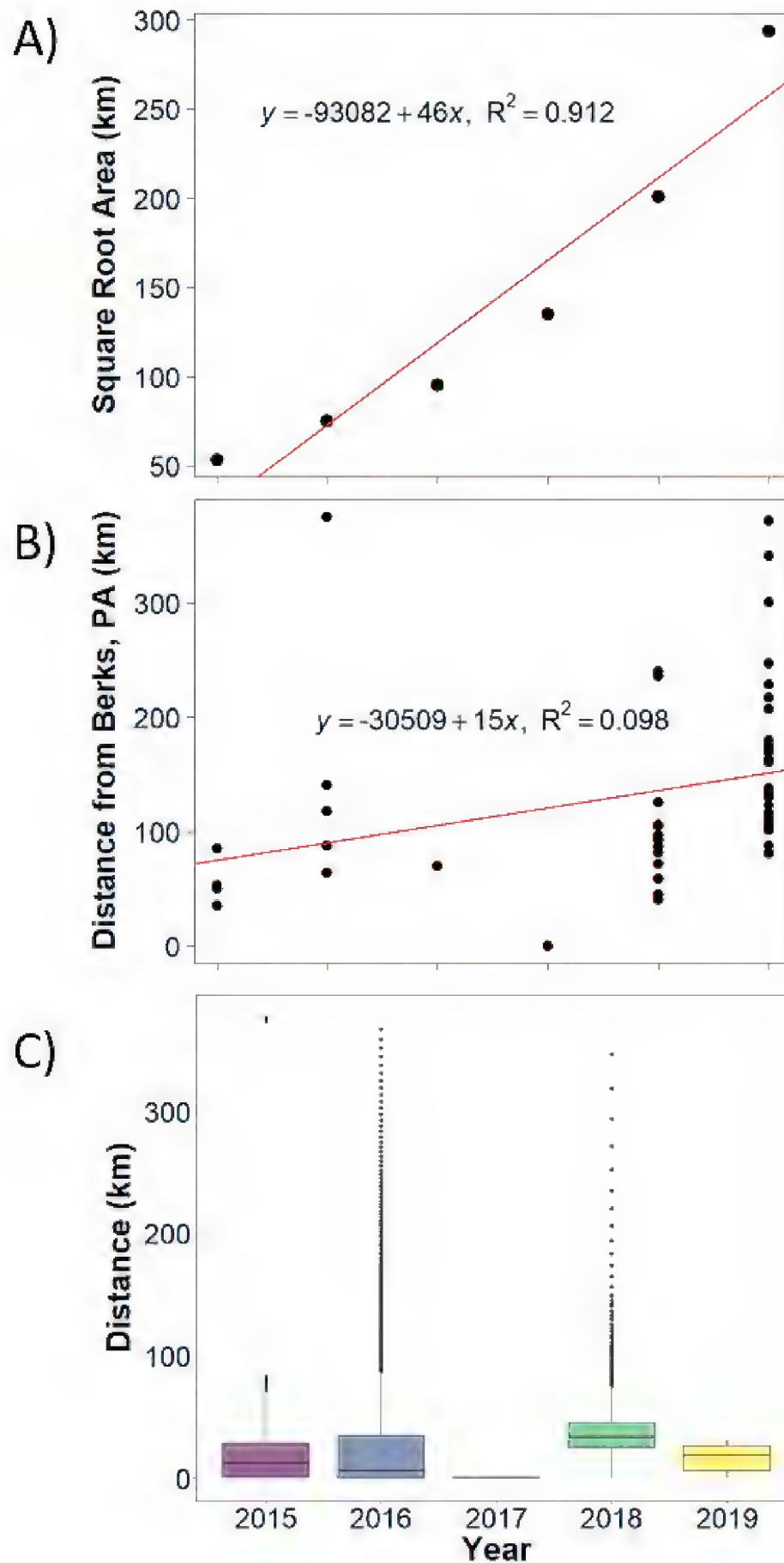
Figure 4. Estimated radial spread rates of spotted lanternfly (SLF) **A** plot of the square root cumulative county area containing SLF establishments divided by π by year of establishment. The slope of the regression is estimated at 46 km per year, providing an estimate of radial spread **B** plot of distance from the centroid of the county with the first SLF detection point (Berks County, PA) by year of establishment. The slope of the regression is estimated at 15 km per year **C** boxplots of boundary displacement distances between years of establishment, with average across all years of 38 km per year and median across all years of 21 km per year.

is also important to note SLF is in the early stages of invasion, and the spread rate may increase as this pest continues to colonize new locations in the USA.

Results of the Cox proportional hazards model suggested that anthropogenic factors, specifically human density, are stronger drivers of SLF spread than forested area or availability of host trees. The role of humans in facilitating spread of invading organisms is a common phenomenon. Known international and domestic pathways of human-mediated spread of tree pests include transportation of pests on live plants (Liebhold 2012) and wood products (e.g. packing materials or movement of firewood) (Yemshanov et al. 2012), though pests can also be transported on non-host materials, such as on stone imports as with SLF. Domestic pathways of human-mediated spread include movement of firewood, transportation via vehicles (e.g. trains, motor vehicles), and "hitchhiking" on travel gear (e.g. hiking gear) and/or pets. Given that SLF lays eggs indiscriminately, human-mediated spread is not limited to host materials. Humans could facilitate the spread of this pest via travel (e.g. automobiles, trains) and movement of both host and non-host materials from an invaded area. Gilbert et al. (2004) came to similar conclusions in their analyses of the horse chestnut leafminer *Cameraria ohridella* Deschka & Dimic (Lepidoptera, Gracillariidae), finding that geographical variation in human population density explained most of the variation in historical spread of this species. Similarly, in an analysis of 79 damaging forest pests, Liebhold et al. (2013) found human population density associated with both spatial proximity and number of invasive forest pests per county across the USA. However, with all such analyses of historical spread, there is always some possibility that statistical associations may be caused in part by more intensive surveying and reporting in more populated areas.

The invasion of tree of heaven in the eastern USA more than 200 years prior to the arrival of SLF may have facilitated the insect's initial establishment, causing an "invasional meltdown" (Feret 1985; Simberloff and Von Holle 1999) in which invasion of one species facilitates the invasion of another. Tree of heaven is the preferred host for SLF and SLF fitness (survival and fecundity) is maximized when feeding on tree of heaven, but this pest can survive and reproduce without access to tree of heaven (Uyi et al. 2020). In addition, SLF's ability to feed on a wide breadth of plant species (more than 70 species) increases the likelihood of the insect encountering a suitable host following dispersal (Dara et al. 2015). The final Cox proportional hazards model did not include a significant effect of tree of heaven abundance on SLF spread, and therefore we found no evidence that this tree species has influenced SLF spread. Surveys for tree of heaven were conducted by many different people including volunteers and residents and thus the resulting data are likely not reflective of true distribution of tree of heaven, despite verification by EDDMapS reviewers. However, given the association of tree of heaven with urban and disturbed environments, it is probable that more accurate tree of heaven distribution data may correlate highly with human population, which emerged as a significant predictor of SLF spread. In addition, as SLF invasion progresses, additional relationships to host trees or other environmental variables may become apparent or the importance of such variables may vary geographically.

Spatial proximity will remain an important predictor in the future spread of this pest, rendering estimation of SLF populations an important step in assessing spread. Current challenges in estimating SLF populations are primarily lack of long-term, systematic population assessment data and difficulties detecting small populations. The SLF-specific dispersal kernel we estimated here provided the best estimates of spatial proximity based on available distribution data but it was limited by the coarse spatial scale of county-level data and the limited temporal replication. We anticipate that as more data are collected on SLF populations, the estimated dispersal kernel could be refined and thus enhance model predictions.

There are a few limitations involved in our study. First, the data used in these analyses consisted of visual surveys that were located based on perceived risk of SLF establishment. These data were not collected in a systematic fashion, and thus there is potential for sampling bias and imperfect detection, e.g. overlooking of individuals. Though work is underway on developing traps to efficiently survey for SLF (Francese et al. 2020), a sensitive SLF-specific trapping system has not yet been widely implemented. The lack of a pest-specific trap increases risk of missed detections in visual assessments, especially for low population densities. Missed detections occur in many invasive species surveillance programs and can lead to underestimation of the extent of species ranges as well as biased estimates of invasion speeds (Guillera-Arroita 2016; Mang et al. 2016). Given these potential biases, we used counties as the unit of analysis, and the estimates of spread rate as well as drivers of local spread at a finer resolution may be different. We also assumed counties with only a single year of SLF detection indicated populations that failed to establish and thus were not detected in future surveys. Failure to establish could be the result of stochastic dynamics or Allee effects, both of which can drive low-density, newly invaded populations to extinction (Liebhold and Tobin 2008). For example, Liebhold and Bascompte (2003) concluded that low density gypsy moth *Lymantria dispar* (L.) populations are likely to reach extinction without intervention, and in their analysis, most of the populations that did go extinct without treatment did so within a year of detection. Where management efforts are in place, failure to establish could also be the result of local eradication efforts. However, there is also a possibility that low-density populations did indeed persist, but due to difficulties in detecting this pest without specific lures or traps, small populations went undetected.

Focusing efforts on assessing populations and on estimating spatial proximity is important in describing and predicting spread of non-native pests. Our findings suggest that SLF has spread from 2014–2020 primarily through local diffusion with less frequent but consistent long-distance dispersal from previously established populations with influence from human populations. Based on the results presented here, we anticipate that SLF will continue to spread in the USA, though management and eradication efforts may effectively reduce population densities, reproductive potential, and ultimately rate of spread. Additional monitoring efforts to prevent and detect long-distance dispersals may prove useful, especially regarding transports of materials from areas with existing SLF populations.

Acknowledgements

References

Barringer LE, Ciafre CM (2020) Worldwide feeding host plants of spotted lanternfly, with significant additions from North America. Environmental Entomology 49(5): 999–1011. https://doi.org/10.1093/ee/nvaa093

Barringer LE, Donovall LR, Spichiger SE, Lynch D, Henry D (2015) The first new world record of *Lycorma delicatula* (Insecta: Hemiptera: Fulgoridae). Entomological News 125: 20–23. 10.3157/021.125.0105.

Bechtold WA, Patterson PL (2005) The enhanced forest inventory and analysis program – national sampling design and estimation procedures. General Technical Report SRS-80. U.S. Department of Agriculture, Forest Service, Southern Research Station, Asheville. https://doi.org/10.2737/SRS-GTR-80

Cornell (2021) New York State Integrated Pest Management: Spotted Lanternfly. https://nysipm.cornell.edu/environment/invasive-species-exotic-pests/spotted-lanternfly/

Dara SK, Barringer L, Arthurs SP (2015) *Lycorma delicatula* (Hemiptera: Fulgoridae): A new invasive pest in the United States. Journal of Integrated Pest Management 6(1): 1–6. https://doi.org/10.1093/jipm/pmv021

Ding J, Wu Y, Zheng H, Fu W, Reardon R, Liu M (2006) Assessing potential biological control of the invasive plant, tree-of-heaven, *Ailanthus altissima*. Biocontrol and Technology 16: 547–566. https://doi.org/10.1080/09583150500531909

EDDMapS (2021) Early Detection & Distribution Mapping System. The University of Georgia Center for Invasive Species and Ecosystem Health. https://www.eddmaps.org/

Eddy D (2018) Spotted Lanternfly Found in Virginia Grapes. Growing Produce. https://www.growingproduce.com/fruits/grapes/spotted-lanternfly-found-virginia-grapes/#:~:text=For%20the%20first%20time%2C%20spotted,in%20Frederick%20County%20in%20January

Edwards J (2021) A Kansas boy entered a unique insect at the state fair. It triggered a federal investigation. The Washington Post. https://www.washingtonpost.com/nation/2021/09/15/kansas-state-fair-spotted-lanternfly/

Fahrner S, Aukema BH (2018) Correlates of spread rates for introduced insects. Global Ecology and Biogeography 27(6): 734–743. https://doi.org/10.1111/geb.12737

Feret PP (1985) Ailanthus: variation, cultivation, and frustration. Journal of Arboriculture 11(12): 361–368.

Francese JA, Cooperband MF, Murman KM, Cannon SL, Booth EG, Devine SM, Wallace MS (2020) Developing traps for the spotted lanternfly, *Lycorma delicatula* (Hemiptera: Fulgoridae). Environmental Entomology 49(2): 269–276. https://doi.org/10.1093/ee/nvz166

Gilbert M, Gregoire JC, Freise JF, Heitland W (2004) Long-distance dispersal and human population density allow the prediction of invasive patterns in the horse chestnut leafminer *Cameraria ohridella*. Journal of Animal Ecology 73(3): 459–68. https://doi.org/10.1111/j.0021-8790.2004.00820.x

Gilbert M, Liebhold, A (2010) Comparing methods for measuring the rate of spread of invading populations. Ecography 33(5): 809–817. https://doi.org/10.1111/j.1600-0587.2009.06018.x

Guillera-Arroita G (2016) Modeling of species distributions, range dynamics and communities under imperfect detection: advances, challenges and opportunities. Ecography 40(2): 281–295. https://doi.org/10.1111/ecog.02445

Hastings A, Cuddington K, Davies KF, Dugaw CJ, Elmendorf S, Freestone A, Harrison S, Holland M, Lambrinos J, Malvadkar U, Melbourne BA, Moore K, Taylor C, Thomson D (2005) The spatial spread of invasions: new developments in theory and evidence. Ecology Letters 8(1): 91–101. https://doi.org/10.1111/j.1461-0248.2004.00687.x

Indiana Department of Natural Resources (DNR) (2021) Spotted lanternfly found in Indiana. DNR Department of Entomology and Plant Pathology.

Jules ES, Kauffman MJ, Ritts WD, Carroll AL (2002) Spread of an invasive pathogen over a variable landscape: A nonnative root rot on Port Orford cedar. Ecology 83(11): 3167–3181. https://doi.org/10.1890/0012-9658(2002)083[3167:SOAIPO]2.0.CO;2

Kovacs KF, Haight RG, McCullough DG, Mercader RJ, Siegert NW, Liebhold AM (2010) Cost of potential emerald ash borer damage in U.S. communities, 2009–2019. Ecological Economics 69(3): 569–578. 10.1016/j.ecolecon.2009.09.004.

Liebhold AM (2012) Forest pest management in a changing world. International Journal of Pest Management 58(3): 289–295. https://doi.org/10.1080/09670874.2012.678405

Liebhold AM, Bascompte J (2003) The Allee effect, stochastic dynamics and the eradication of alien species. Ecology Letters 6: 133–140. 10.1046/j.1461-0248.2003.00405.x.

Liebhold AM, Tobin PC (2008) Population ecology of insect invasions and their management. Annual Review of Entomology 53: 387–408. https://doi.org/10.1146/annurev.ento.52.110405.091401

Liebhold AM, Kean JM (2019) Eradication and containment of non-native forest insects: successes and failures. Journal of Pest Science 92(1): 83–91. https://doi.org/10.1007/s10340-018-1056-z

Liebhold AM, Brocherhoff EG, Garrett LJ, Parke JL, Britton KO (2012) Live plant imports: the major pathway for forest insect and pathogen invasions of the US. Frontiers in Ecology and the Environment. 10: 135–143. https://doi.org/10.1890/110198

Liebhold A, McCullough D, Blackburn L, Frankel S, Von Holle B, Aukema J (2013) A highly aggregated geographical distribution of forest pest invasions in the USA. Diversity and Distributions 19(9): 1208–1216. https://doi.org/10.1111/ddi.12112

Mang T, Essl F, Moser D, Karrer G, Kleinbauer I, Dullinger S (2016) Accounting for imperfect observation and estimating true species distributions in modelling biological invasions. Ecography 40(10): 1187–1197. https://doi.org/10.1111/ecog.02194

Orlova-Bienkowskaja MJ, Bienkowski AO (2018) Modeling long-distance dispersal of emerald ash borer in European Russia and prognosis of spread of this pest to neighboring countries within next 5 years. Ecology and Evolution 8(18): 9295–9304. https://doi.org/10.1002/ece3.4437

Parra G, Moylett H, Bulluck R (2017) Technical working group summary report: spotted lanternfly, *Lycorma delicatula* (White, 1845). USDA APHIS Plant Protection and Quarantine Center for Plant Health Science and Technology. https://www.agriculture.pa.gov/Plants_Land_Water/PlantIndustry/Entomology/spotted_lanternfly/research/Documents/SLF%20TWG%20Report%20020718%20final.pdf

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Scheid L (2020) Spotted lanternfly egg masses found in Maine were traced to Pennsylvania. Reading Eagle. https://www.readingeagle.com/news/spotted-lanternfly-egg-masses-found-in-maine-were-traced-to-pennsylvania/article_85071664-12e0-11eb-837e-a7e-0cd5a2f28.html

Sharov AA, Liebhold AM (1998). Bioeconomics of managing the spread of exotic species with barrier zones. Ecological applications, 8: 833–845. 10.1111/j.0272-4332.2004.00486.x.

Simberloff D, Von Holle B (1999) Positive interactions of nonindigenous species: Invasional meltdown? Biological Invasions 1: 21–32. https://doi.org/10.1023/A:1010086329619

Skellam JG (1951) Random dispersal in theoretical populations. Biometrika 38: 196–218. https://doi.org/10.1093/biomet/38.1-2.196

Shigesada N, Kawasaki K, Takeda Y (1995) Modeling stratified diffusion in biological invasions. The American Naturalist 146: 229–251. https://doi.org/10.1086/285796

Thomas L, Reyes EM (2014) Tutorial: Survival estimation for Cox regression models with time-varying coefficients using SAS and R. Journal of Statistical Software 61(CS1): 1–23. https://doi.org/10.18637/jss.v061.c01

Tobin PC, Liebhold AM, Roberts EA (2006) Comparison of methods for estimating the spread of a non-indigenous species. Journal of Biogeography 34(2): 305–312. 10.1111/j.1365-2699.2006.01600.x

Urban JM (2019) Perspective: Shedding light on spotted lanternfly impacts in the USA. Pest Management Science 76(1): 10–17. https://doi.org/10.1002/ps.5619

Urban JM, Calvin D, Hills-Stevenson J (2021) Early response (2018–2020) to the threat of spotted lanternfly, *Lycorma delicatula* (Hemiptera: Fulgoridae) in Pennsylvania. Annals of the Entomological Society of America. https://doi.org/10.1093/aesa/saab030

USDA Animal and Plant Health Inspection Services (USDA APHIS) (2018) Spotted lanternfly eradication program in select counties of Pennsylvania, Supplemental Environmental Assessment. U. S. Department of Agriculture, Animal and Plant Health Inspection Service, Plant Protection and Quarantine, Riverdale, MD. https://www.aphis.usda.gov/plant_health/ea/downloads/2018/slf-pa-supplemental-ea.pdf

USDA Forest Service (2016) USFS Percent Tree Canopy (Cartographic Version). Edition 1.0. Salt Lake City, UT. https://data.fs.usda.gov/geodata/rastergateway/treecanopycover/

USDA Forest Service (2019) Forest Inventory and Analysis National Core Field Guide. Volume I: Field Data Collection Procedures for Phase 2 Plots, Version 9.0. https://www.fia.fs.fed.us/library/field-guides-methods-proc/docs/2019/core_ver9-0_10_2019_final_rev_2_10_2020.pdf

Uyi O, Keller JA, Johnson A, Long D, Walsh B, Hoover K (2020) Spotted lanternfly (Hemiptera: Fulgoridae) can complete development and reproduce without access to the preferred host, *Ailanthus altissima*. Environmental Entomology 49(5): 1185–1190. https://doi.org/10.1093/ee/nvaa083

Wakie TT, Neven LG, Yee WL, Lu Z (2020) The establishment risk of *Lycorma delicatula* (Hemiptera: Fulgoridae) in the United States and globally. Journal of Economic Entomology 113: 306–314. https://doi.org/10.1093/jee/toz259

Ward SF, Fei S, Liebhold AM (2020) Temporal dynamics and drivers of landscape-level spread by emerald ash borer. Journal of Applied Ecology 57: 1020–1030. https://doi.org/10.1111/1365-2664.13613

Yemshanov D, Koch FH, Ducey M, Koehler K (2012) Trade-associated pathways of alien forest insect entries in Canada. Biological Invasions 14: 797812. https://doi.org/10.1007/s10530-011-0117-5

Supplementary material 1

Figure S1

Authors: Rachel T. Cook

Data type: map

Explanation note: Locations of SLF visual syrveys conducted by the US Animal and Plant Health Inspection Service and Pennsylvania Department of Agriculture.

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/neobiota.70.67950.suppl1

Supplementary material 2

Figure S2. Distributions of anthropogenic predictor variables used in Cox proportional hazards moder development

Authors: Rachel T. Cook

Data type: maps

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/neobiota.70.67950.suppl2

Supplementary material 3

Figure S3. Habitat predictor variable distributions

Authors: Rachel T. Cook

Data type: maps

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/neobiota.70.67950.suppl3

Supplementary material 4

Table S1. Predictor variable summary

Authors: Rachel T. Cook

Data type: statistical data

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/neobiota.70.67950.suppl4